УДК 004.051:004.627

В.І. Бовсунівський
Інститут кібернетики ім. В.М. Глушкова НАН України, м. Київ, Україна
vbi@ukr.net

Доцільність реалізації оперативно-адаптивного стиску зображень та відеоданих

У статті розглядаються основні методи кодування відеоданих та їх недоліки. Розглядається проблема спотворення форми обвідних відеосигналів за рахунок кодування при дослідженні тестових кадрів первинних масивів відеоданих.

Вступ

За рахунок розповсюдження таких технологій, як DVD-відео, цифрового телебачення та потокового відео в мережі Інтернет, вагомим компонентом їх реалізації стала компресія відеоданих. Успіх цифрового телебачення та DVD-відео базується на технології MPEG-2, який був розроблений 15 років тому. Ця технологія довела свою ефективність, проте сьогодні вона є застарілою. На її зміну прийшли нові засоби стиснення даних, такі як MPEG-4 та H.264, які показали хороші результати в плані стиску та якості відео. Алгоритми, що реалізують ці методи, використовують ту ж саму схему кодування й декодування, що і MPEG-2, відмінність їх роботи полягає у вдосконаленні основних етапів реалізації даної схеми. Проте, як і раніше, ці технології використовують стиск із втратами даних, що в результаті так чи інакше якісно впливає на сприйняття людиною такого відео [1].

Метою даної роботи є обґрунтування основних моментів втрати корисної інформації алгоритмом, який реалізований тим чи іншим методом кодування відеоданих, а також візуальна та аналітична оцінка якості кодування відеоданих при різних коефіцієнтах стиску.

Основні принципи стиску відеоданих та їх недоліки

Для стиснення відеосигналу використовується видалення різного роду збитковостей. Зокрема, статистична збитковість видаляється таким чином, що вихідний сигнал може бути повністю відновлений за допомогою декодера практично без втрат даних. Проте методи стиску даних без втрат дають досить скромні результати, тому будь-який алгоритм стиску відеоданих не обходиться лише цими методами. Окрім видалення статистичної збитковості алгоритми відеокомпресії видаляють збитковість з часової, просторової та частотної областей сигналу.

Сам відеокодер складається з трьох основних компонент: часової моделі, просторової моделі та ентропійного кодера. На вхід часової моделі подається нестиснений цифровий відеосигнал. Часова модель зменшує часову збитковість, використовуючи високий рівень кореляції між послідовними відеокадрами. Зазвичай така модель будує прогноз для наступних кадрів на основі сусідніх кадрів послідовності. За рахунок цього зменшується час пошуку часових збитковостей. Прогноз формується за одним чи кіль-

кома попередніми кадрами, при цьому здійснюється корекція розходження кадрів. Очевидно, що за рахунок похибок округлення при кодуванні та відновленні, а також через інші причини, така послідовність залишкових кадрів втрачатиме корисні дані. Тому, зазвичай, така модель використовує ключові кадри, які зберігаються повністю. Такі кадри вставляються кожні 15 – 20 кадрів для контролю втрати даних.

На виході часової моделі ми отримуємо залишковий кадр та деяку множину числових параметрів моделі. Залишковий кадр є входом для просторової моделі, яка використовує подібність сусідніх просторових фрагментів даного кадру, усуваючи цим самим просторову збитковість. Це реалізується застосуванням ряду перетворень до залишкового кадру та квантуванням результату. Застосування перетворень до залишкового кадру переводить його до іншої області, в якій він набуває форму послідовності коефіцієнтів використовуваного перетворення. Ці коефіцієнти квантуються методом видалення несуттєвих значень, в результаті чого залишається невелика кількість суттєвих коефіцієнтів, які дають більш компактну форму залишкового кадру. Виходом просторової моделі є множина квантованих коефіцієнтів перетворення.

Параметри часової та просторової моделі стискаються ентропійним кодером. При цьому видаляється статистична збитковість. У результаті цього отримуються компактні дані, які можна передавати по мережах чи зберігати на цифрових носіях. Відеодекодер реконструює відеокадр по стисненому бітовому потоку даних. Коефіцієнти та вектори руху декодуються ентропійним декодером, після чого просторова модель декодує отримані дані і формує частину залишкового кадру. Декодер використовує параметри векторів руху разом з одним чи кількома раніше декодованими кадрами для побудови прогнозу поточного кадру, а сам кадр реконструюється додаванням залишкового кадру до цього кадру-прогнозу.

Зазвичай різниці між кадрами можуть виникати через рух об’єктів на відео, які не міняють свою форму (рух автомобіля, рух людини, панорамне відео, та ін.). Рух таких об’єктів на відео зазвичай ініціюють переміщення цілих груп пікселів між кадрами, які в цілому утворюють оптичний потік. Якщо точно знати оптичний потік пікселів відео, то можна здійснити точний прогноз зміщення більшості пікселів на сусідніх кадрах. Проте точне обчислення оптичного потоку не є ефективним через високу складність обчислень. Тому на практиці широко використовується метод компенсації руху, який компенсує переміщення прямокутних областей. Такий метод включає наступні етапи:

1. Пошук на попередньому (наступному) кадрі блоку пікселів, для яких можна застосувати цей метод.
2. Отримання залишкового блока.
3. Кодування залишкового блока

Даний метод достатньо простий в реалізації. Прямокутні області досить просто вписуються в прямокутний формат кадрів, а також в процедури перетворення зображень на основі блоків. Проте реальні об’єкти на відео зазвичай мають неправильну форму, в результаті чого виникають деякі труднощі обробки країв таких об’єктів даним методом. Крім того об’єкти часто зміщуються на неціле число пікселів між кадрами, що спричиняє певну втрату даних. Зазвичай ці недоліки спричиняють втрату кореляції між окремими такими блоками, що часто добре видно на відео у вигляді блочної структури кадру [1].

Якість кодування відеоданих

Оцінка якості відеоданих є дуже складним процесом. Зокрема, характер оцінювання залежить від результату, який очікується від відео. Виділяють суб’єктивні та об’єктивні методи оцінювання якості відеоданих. Візуальне сприйняття відеоданих людиною є суб’єктивним від самого початку, оскільки на нього діє ряд факторів, як внутрішніх

так і зовнішніх, в залежності від яких відео буде сприйнято людиною по-різному. Тому під таким оцінюванням слід вважати якість показу даного відео, а не якість самого відео. Для отримання більш точних результатів проводяться статистичні експерименти, де на основі вражень від перегляду великої кількості неспокушених глядачів отримується статистична інформація.

До об'єктивних методів оцінки якості відеоданих відносять автоматичні методи і алгоритми вимірювання якості. Найбільш популярним методом оцінки якості відео є пікове співвідношення сигнал/шум PSNR, що оцінює чистоту та достовірність сигналу, якості алгоритмів стиску відео, яке призначене не для відеопоказу, а для обробки даних іншими спеціалізованими алгоритмами (обробка знімку флюорографії, дані космічних знімків, відеоспостереження в охоронних системах тощо.)

Ми провели об'єктивну оцінку якості кодування відео в форматах MPEG2 та H.264 для різних коефіцієнтів стиску. Для цього ми представили відеодані у вигляді суцільного потоку пікселів, розкладених на відліки червоного, зеленого і синього кольорів. Потік кожного відліку за рахунок кореляції сусідніх пікселів утворював криву. Таку криву відео без стиску ми порівняли шляхом накладання з кривою відео зі стиском різного ступеня (рис. 1), в результаті чого отримали, що при кодуванні MPEG2 та H.264 навіть при незначному збільшенні коефіцієнтів стиску починається втрата корисних даних.

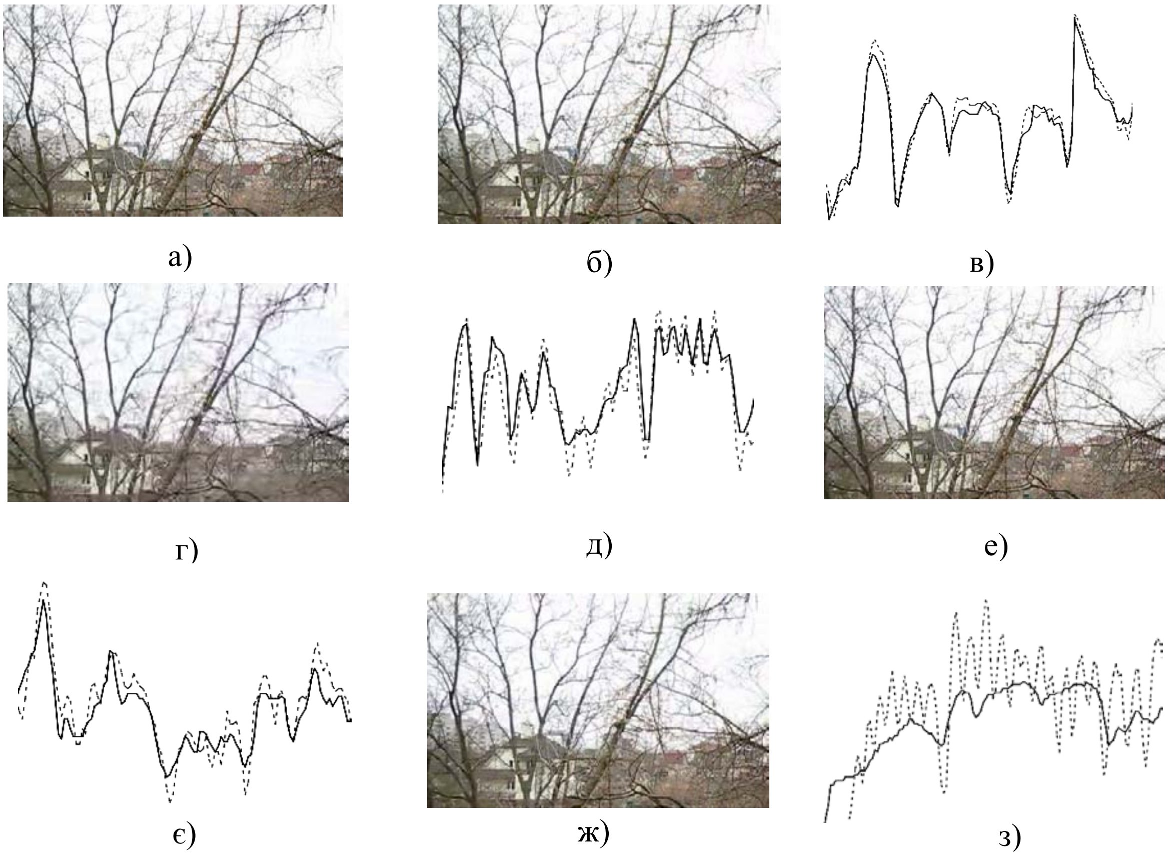


Рисунок 1 – Порівняння кривих червоного відліку контрольного зразку і перекодованого відео (пунктиром – оригінальна крива; суцільною – крива перекодованого відео): а) контрольний зразок; б) кадр з середнім коефіцієнтом стиску MPEG2; в) порівняння відліків кадру б) з контрольним зразком; г) кадр з високим коефіцієнтом стиску MPEG2; д) порівняння відліків кадру г) з контрольним зразком; е) кадр з середнім коефіцієнтом стиску H.264; є) порівняння відліків кадру е) з контрольним зразком; ж) кадр з високим коефіцієнтом стиску H.264; з) порівняння відліків кадру ж) з контрольним зразком

На такому порівнянні крива перекодованого відео починає відхилятися від тенденції кривої оригінального відео на різких і високих стрибках амплітуди кривої. Причому крива після кодування MPEG2 набуває шумів технічного походження, а крива після кодування H.264 виглядає більш розгладженою. За рахунок цього при збільшенні коефіцієнтів стиску відео, перекодоване за допомогою MPEG2, набуває артефактів, які спотворюють візуальні дані, а відео, перекодоване за допомогою H.264, стає більш розмитим, що візуально не помітно, проте таке розмиття спричиняє втрату дрібних деталей [2]. Саме це робить непридатними такі методи кодування, як MPEG2 та H.264, для кодування спеціалізованого відео. Зазвичай на практиці таке відео не стискається, а зберігається необробленим, що вимагає великої кількості ресурсів та дорогого обладнання.

Висновки

Сам факт використання ключових кадрів свідчить про те, що метод видалення часової збитковості може породжувати спотворення відео, і чим рідше використовуються ключові кадри в потоці, тим більше спотворень набуватиме відео.

Практично найбільше спотворень спричиняє метод компенсації руху, який компенсує рух прямокутних областей. Цей метод використовує компенсацію переміщення прямокутних областей пікселів. На практиці зазвичай форми мають неправильну форму, тому виникають труднощі при використанні цього методу на краях таких областей. Зазвичай існує проблема кореляції таких областей на відновленому кадрі. На кадрі після стиску H.264 грань між такими областями розмивається, проте це спричиняє втрату дрібних деталей. Крім того втрату даних спричиняє зміщення блоку на неціле число пікселів. І навіть усереднення сусідніх значень пікселів спричиняє певні втрати інформації.

Література

1. Ричардсон Я. Видеокодирование. H.264 и MPEG-4 – стандарты нового поколения / Ричардсон Я. – М. : Техносфера, 2005. – 365 с.
2. Бовсунівський В.І. Дослідження методів і алгоритмів оперативного відновлення обвідних відеосигналів / В.І. Бовсунівський // Праці проблемно-наукової міжгалузевої конференції «Інформаційні проблеми комп’ютерних систем, юриспруденції, економіки та моделювання». – Бучач, 2010 (Матеріали знаходяться у друці).

В.И. Бовсуновский

Целесообразность реализации оперативно-адаптивного сжатия изображений и видеоданных

В статье рассматриваются основные методы кодирования видеоданных и их недостатки. Рассматривается проблема искажения формы обводящих видеосигналов за счет кодирования при исследовании тестовых кадров первичных массивов видеоданных.

V.I. Bovsunivskiy

Feasibility of Operational and Adaptive Image and Video Data Compression

This article discusses the basic methods of encoding video data and their shortcomings. It describes the problem of distortion waveforms of video when encoding in the study of test frames of the primary array of video data.

Стаття надійшла до редакції 28.05.2010.